

NSTEPS Data Analysis Support and Data Consultation Exercise

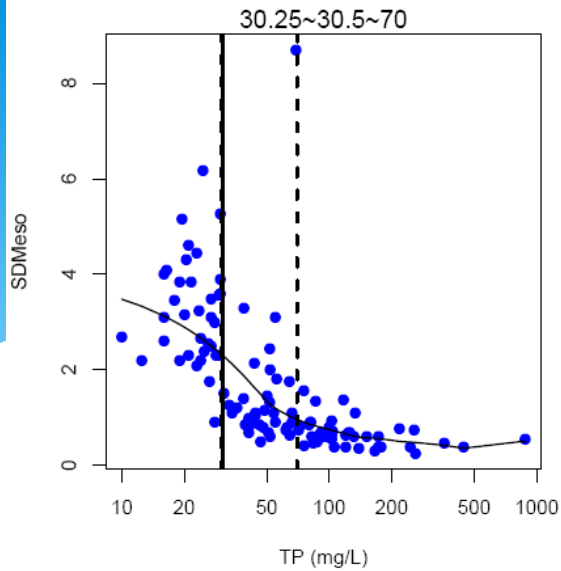
Outline

- * General overview of NSTEPS analysis support process
- * Sample data analysis project
- * Sample data analysis output
- * General Discussion

NSTEPS Data Analysis Support

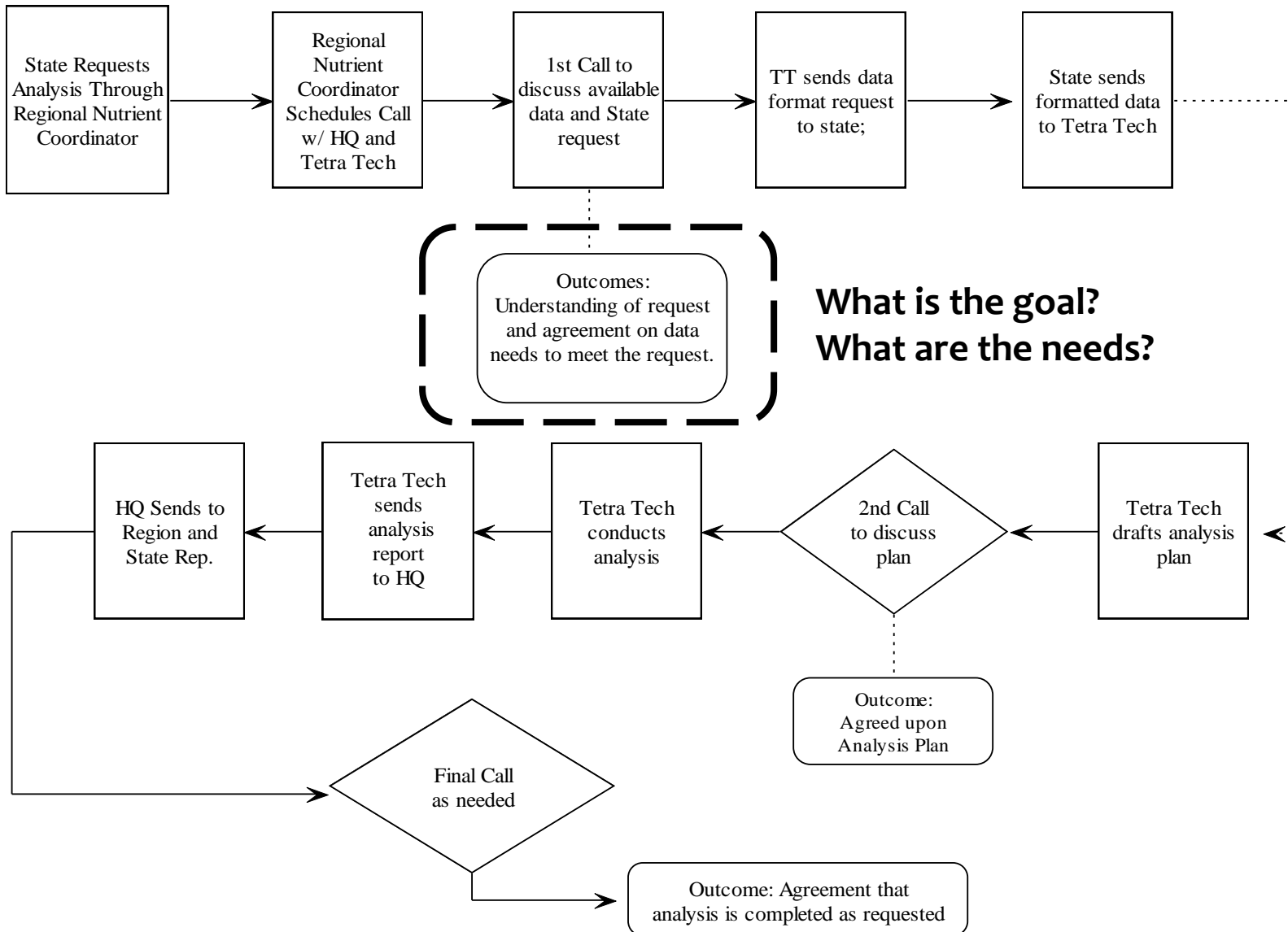
- * Provide analytical support for state/tribal numeric nutrient criteria derivation
- * Collaboration between state/tribe and EPA NSTEPS
- * Not intended to constitute numeric nutrient criteria derivation per se.

Example Activities

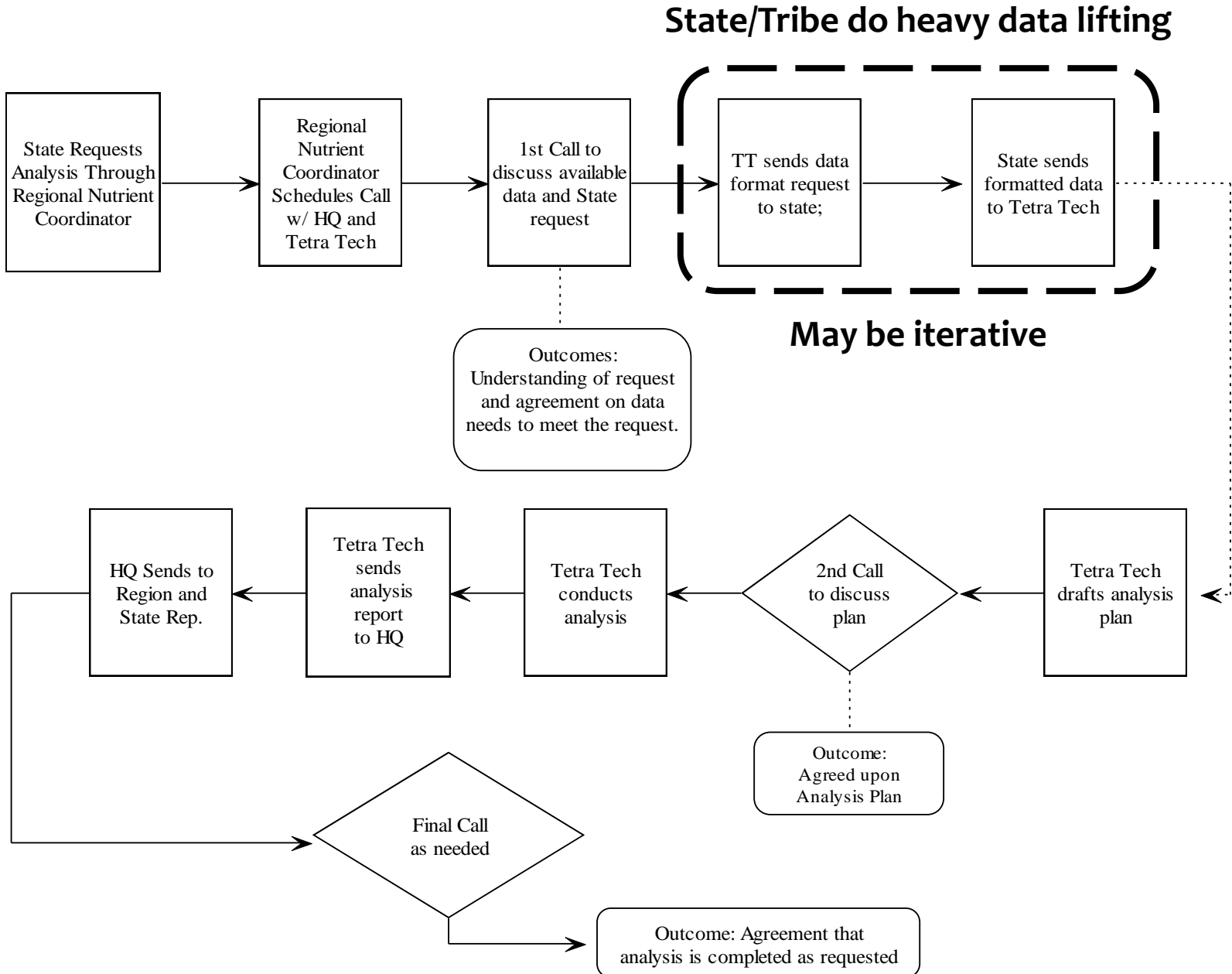


- * Conduct waterbody classifications
- * Derive regional reference distribution values
- * Provide stressor-response analyses in support of stream nutrient criteria derivation
- * Provide a QA of analyses conducted
- * Provide supporting scientific literature for analyses

PROCESS FOR N-STEPS NUTRIENT DATA ANALYSIS



PROCESS FOR N-STEPS NUTRIENT DATA ANALYSIS



Data Transfer Guidelines

- * How to submit data to NSTEPS
- * Includes:
 - * Specific instructions
 - * Checklist...

Data Transfer Guidelines

Checklist			
Item	Description	Check	Notes
Non-numeric values	Remove any non-numeric values from numeric value fields. Account for those with any meaning by transforming or screening the value as applicable.	<input type="checkbox"/>	
Outliers Screened?	Check and verify outliers. Report in the Notes or elsewhere how screening was performed.	<input type="checkbox"/>	
Units Indicated	Units for each parameter included are specified.	<input type="checkbox"/>	
Censored Data	Address censored data (e.g. values reported as "<" or "0"). Include an explanation of rules/procedures to document or process censored values.	<input type="checkbox"/>	
Zero or Negative values	Eliminate zero or negative values (unless legitimate, e.g., temperature).	<input type="checkbox"/>	
Error Check	100% error check the data files for transposition errors and typos. Check for errors in related variables (e.g. SRP>TP, or NH3N>TKN)	<input type="checkbox"/>	

Data Transfer Guidelines

- * How to submit data to NSTEPS
- * Includes:
 - * Specific instructions
 - * Checklist
 - * Data Tables
 - * Station attributes
 - * Chemistry
 - * Biology/Response

Data Transfer Guidelines

StationID	Station Name	Lat	Long	Waterbody Identifier	Waterbody Name	Waterbody Type	Waterbody Area/length (units)	Other Attribute ₁
EXAMPLESTA001	Franklin Lake North End	40.482853	-77.357032	2465A	Franklin Lake	Lake	75	
EXAMPLESTA002	Franklin Lake South End	40.482953	-77.356751	2465A	Franklin Lake	Lake	75	
EXAMPLESTA003	Lake Martin	40.496432	-77.356027	2232	Lake Martin	Lake	12	
EXAMPLESTA004	Thornton Reservoir Intake	40.465932	-77.352319	1259C	Thorton Reservoir	Lake	32	
EXAMPLESTA005	Thorton Reservoir Deep	40.465802	-77.3521	1259C	Thorton Reservoir	Lake	32	

StationID	Date	Time	Depth (units)	Depth	TN (units)	TP (units)	Chl-a C (units)	DO (units)	SD (units)	TempC (units)	Comments	Other Parameter ₁
EXAMPLESTA001	3/15/2000	700	0.5	Surface	3.23	0.052	12.22	3.61		23.4		
EXAMPLESTA001	3/15/2000	1550	0.5	Surface	3.32	0.061	14.3	3.72		23.6		
EXAMPLESTA001	3/15/2000	1550	5.1	Bottom				2.53	4.2	21.5		
EXAMPLESTA001	6/12/2000	820	0.5	Surface	3.41	0.056	14.65	3.82		26.3		
EXAMPLESTA002	4/5/2001	832	0.5	Surface	5.23	0.067	20.23	3.45		23.5		

StationID	Date	Time	Method	Repl icate	Total Taxa	Total Abundance	Other Attribute ₁	Other Attribute ₂	...	Other Attribute _x	Taxon 1	Taxon 2
EXAMPLESTA001	3/15/2000	700	EMAP	1	3	12					6	
EXAMPLESTA002	4/5/2001	832	EMAP	1	4	31					2	3
EXAMPLESTA002	8/5/2003	1105	EMAP	1	3	19					4	
EXAMPLESTA003	6/22/2004	1000	EMAP	1	2	27						3

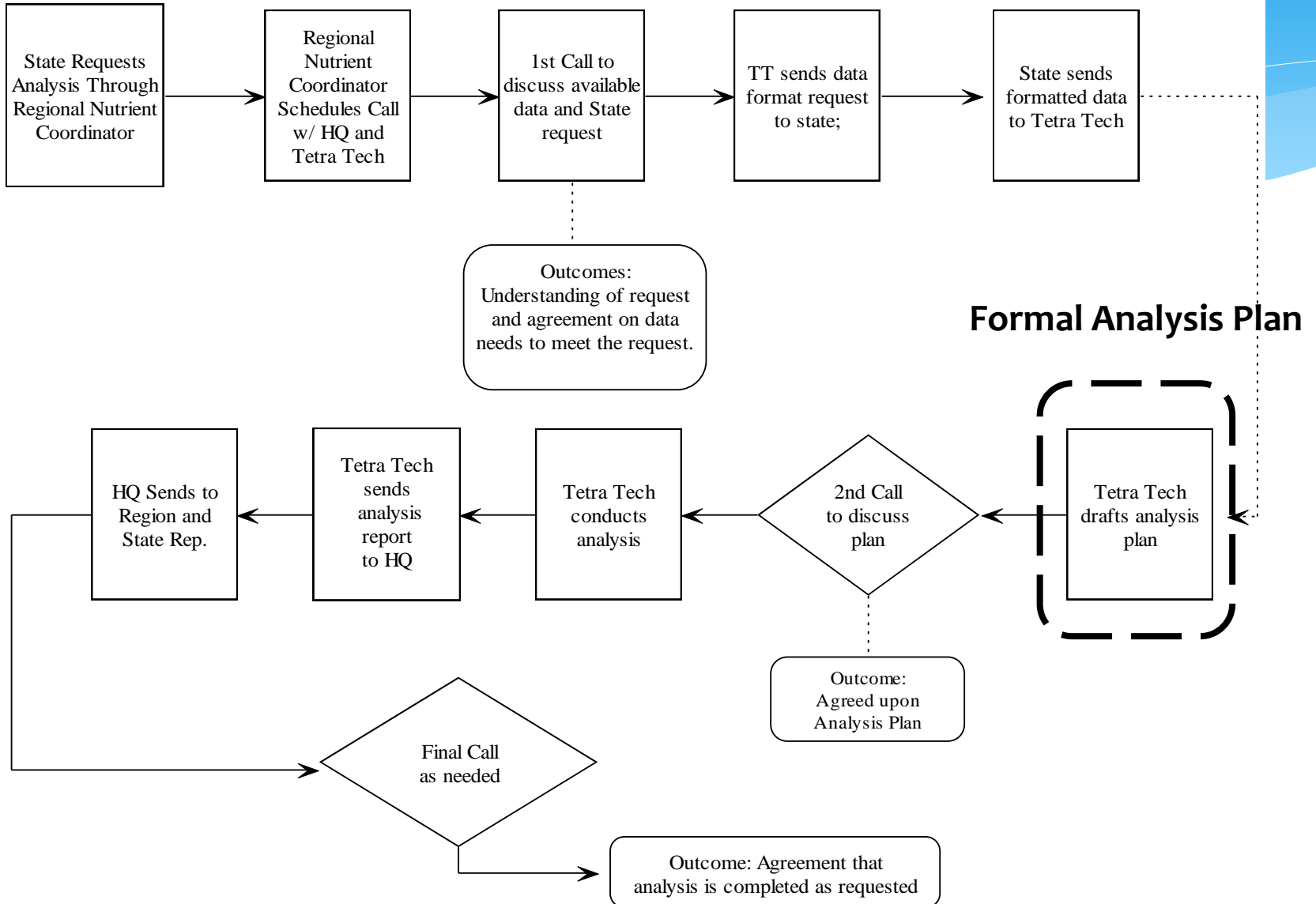
Data Transfer Guidelines

- * How to submit data to NSTEPS
- * Includes:
 - * Specific instructions
 - * Checklist
 - * Data Tables
 - * Station attributes
 - * Chemistry
 - * Biology/Response

Data Review

- * NSTEPS may conduct a data review and submit any questions back to state
- * Iterative process to achieve final appropriate dataset
- * State/tribe, again, bears responsibility for data preparation
- * A Data QA Report would be especially helpful (see example)

PROCESS FOR N-STEPS NUTRIENT DATA ANALYSIS



Data Analysis Plan

For

**Analysis of Idaho Nutrient and Biological Data for the
Nutrient Scientific Technical Exchange Partnership Support
(N-STEPS)**

Analysis Plan

- * Objectives, Introduction and Analysis Goals
 - * Sample:
 - * Can visible slime growths or nuisance aquatic growths be defined quantitatively?
 - * Are nutrients associated with these growths in a stressor-response context?
 - * Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?
- * Methods
 - * Sample: Distribution analysis, Modeled Reference, Stressor-Response

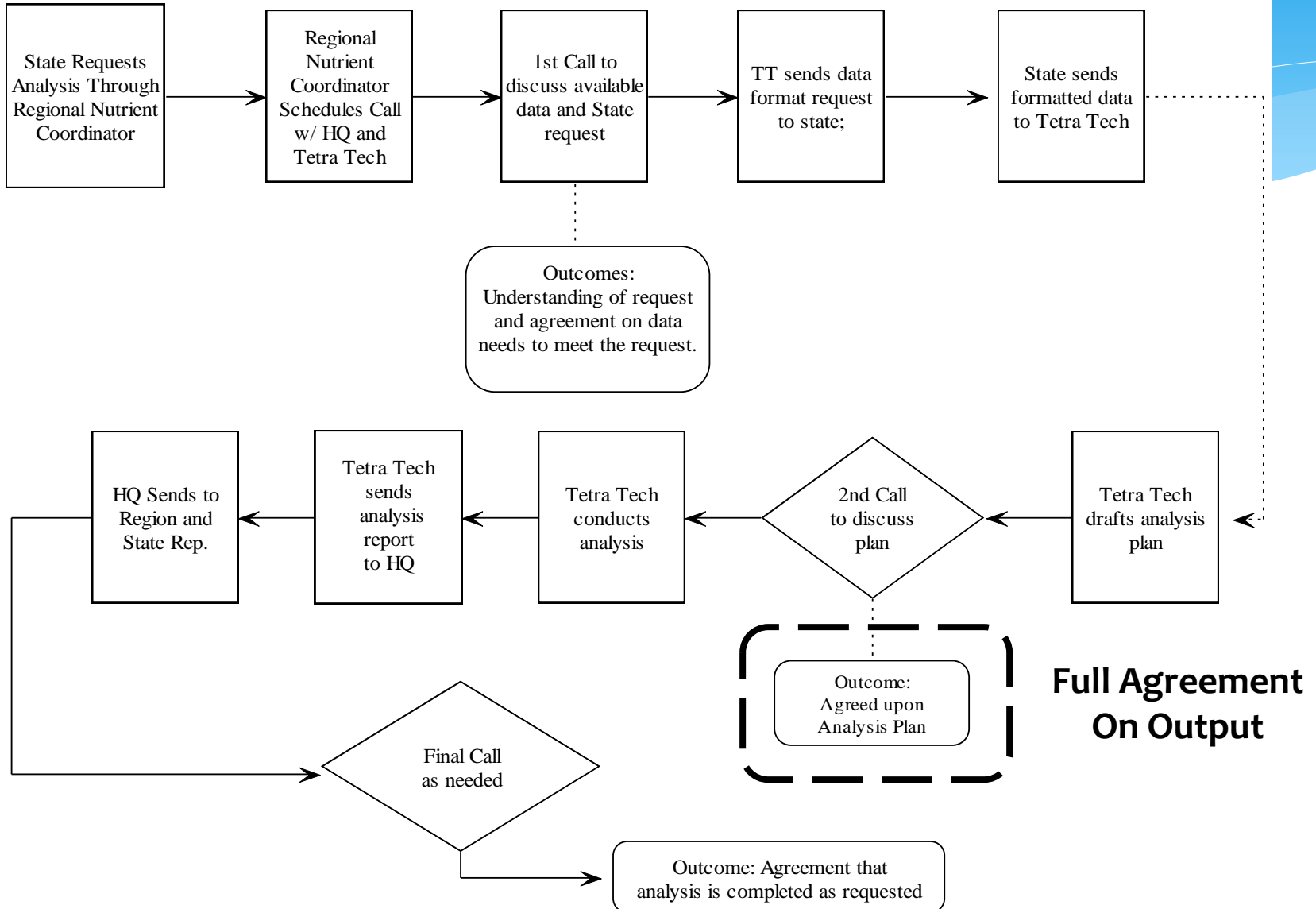
Analysis Plan

- * Outcomes

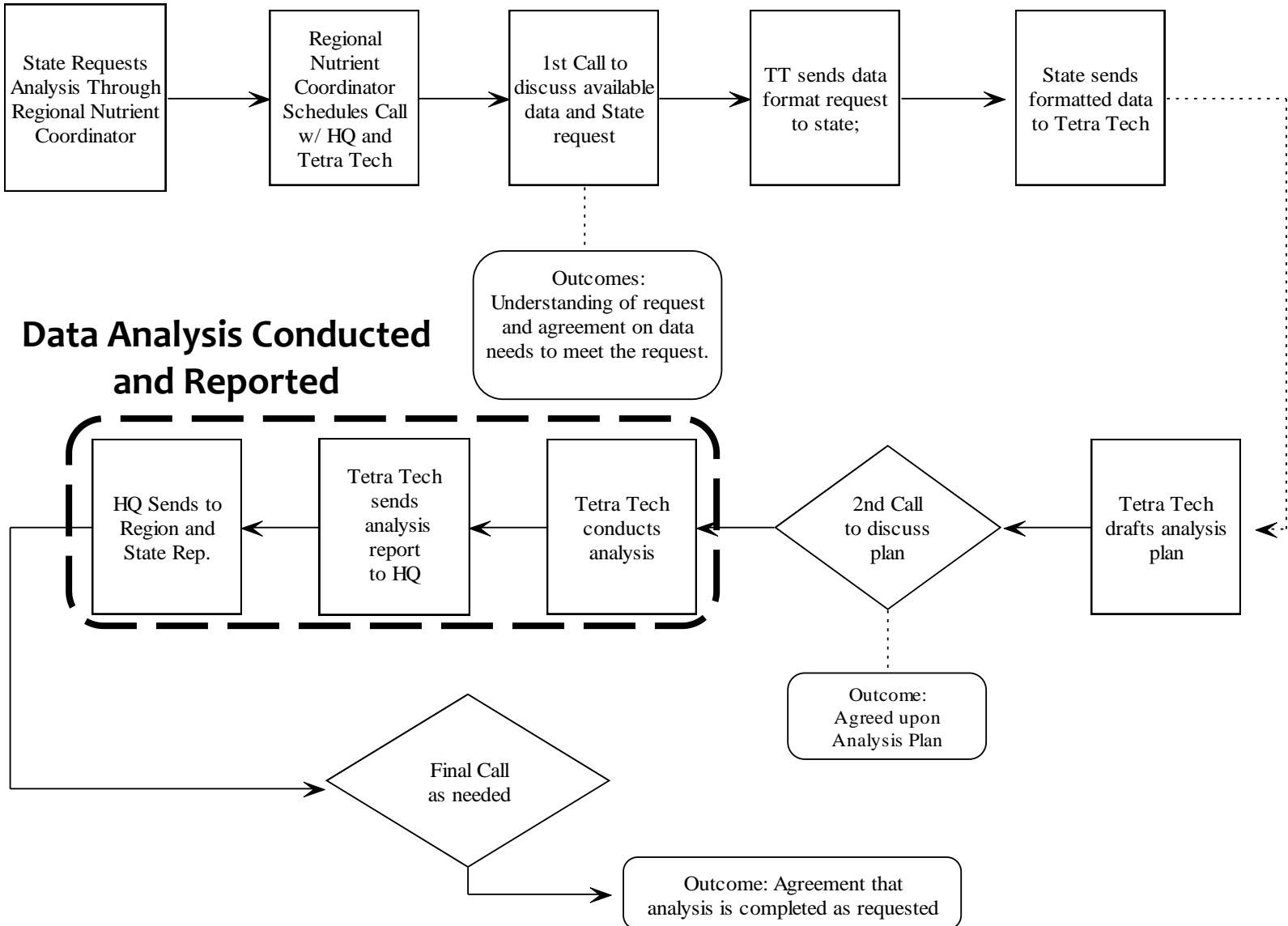
- * Sample:

- * Site classes and co-varying environmental variables used to reduce natural variability in the nutrient data.
 - * Nutrient and response endpoints determined from frequency distribution analysis by class
 - * Nutrient endpoints determined from modeled reference expectation including the regression equations, regression model diagnostics, and the estimated values.
 - * Nutrient endpoints from stressor response analysis including visual plots of interest, linear regression and LOWESS curve fits, interpolated endpoints, and thresholds determined using visual estimates with LOWESS and change-point analysis, if applicable.
 - * Uncertainty estimates for each analysis.

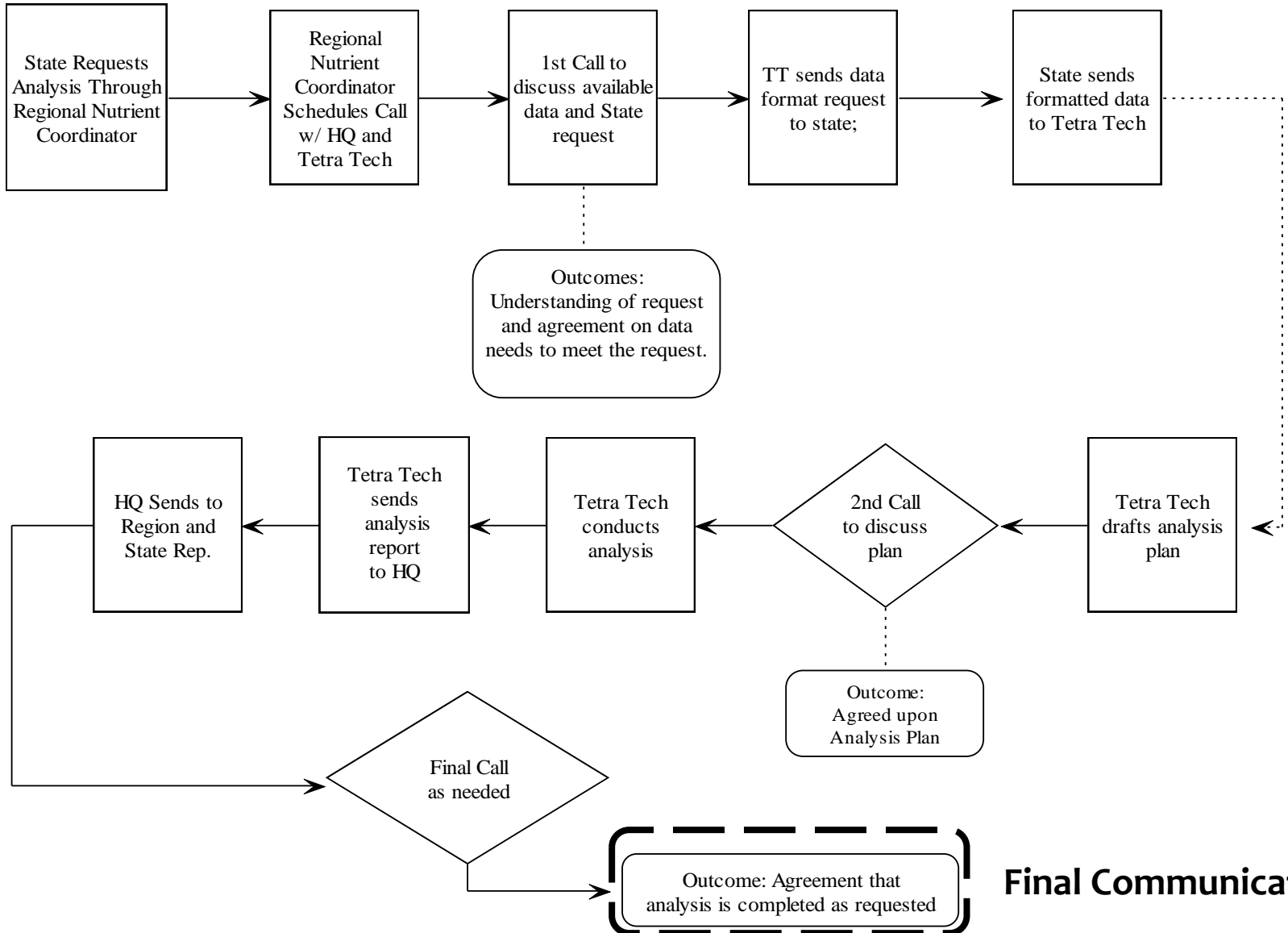
PROCESS FOR N-STEPS NUTRIENT DATA ANALYSIS



PROCESS FOR N-STEPS NUTRIENT DATA ANALYSIS



PROCESS FOR N-STEPS NUTRIENT DATA ANALYSIS



DATA CONSULTATION

EXERCISE: IDAHO

Goal

- How might an analysis plan develop and what types of analyses could be run?
- Questions are a great start (ideally before data collection)
 - ▣ Idaho questions as an example
- Springboard for additional analysis ideas/discussion

Idaho's questions

- Can visible slime growths or nuisance aquatic growths be defined quantitatively?
- Are nutrients associated with these growths in a stressor-response context?
- Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?
- Data gaps –
 - ▣ e.g., duration, frequency or magnitude that can be associated with nutrient levels?
 - Best bang for the buck in sampling strategy?

Can “visible slime growths” or “nuisance aquatic growths” be defined quantitatively?

□ Observations

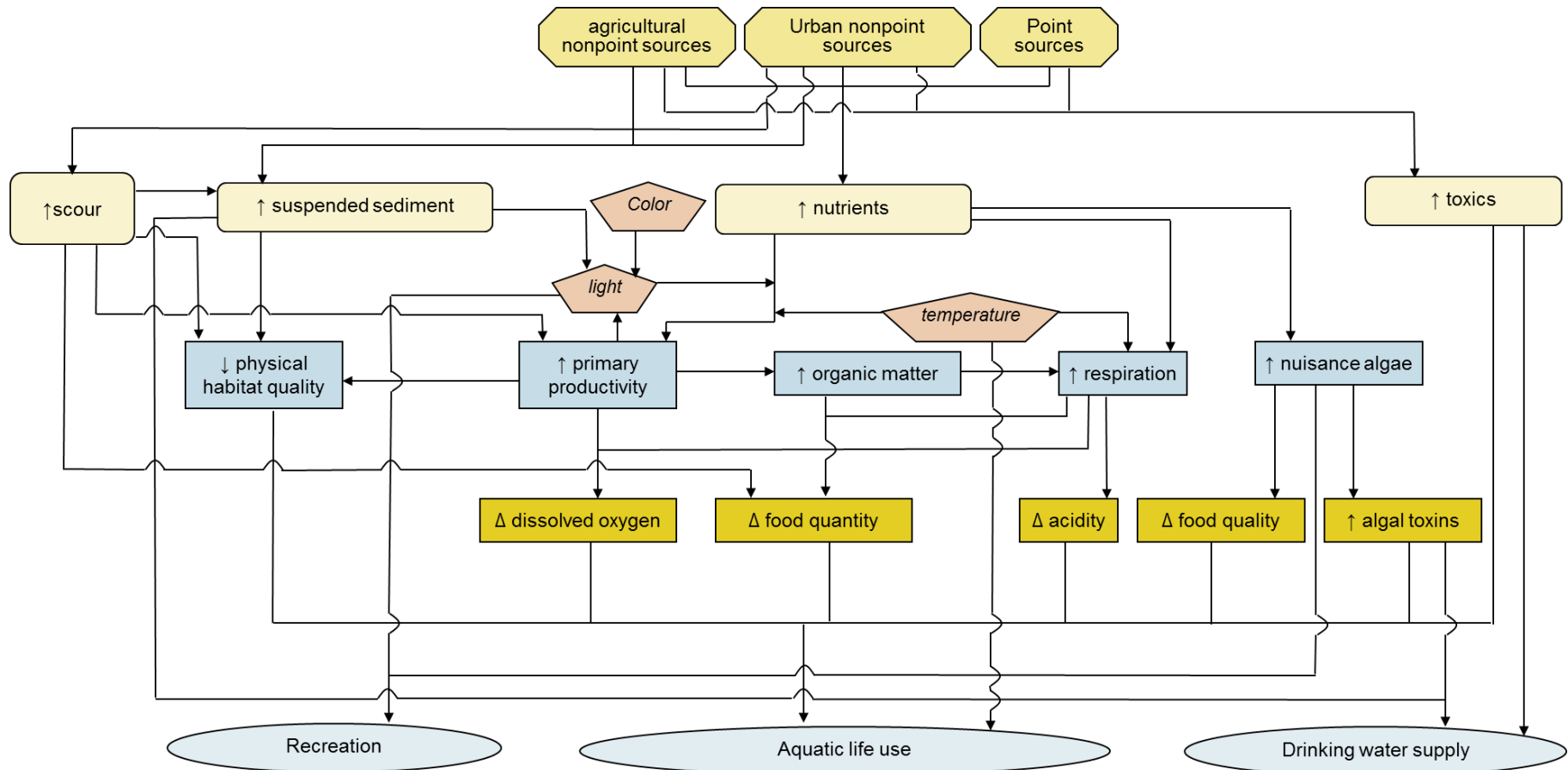
- Algal effects to recreational use ratings
- Algal aesthetic ratings
- Percent coverage of:
 - green filamentous algae
 - floating mats and scum
 - blue-green algae, diatom mats, red algae
 - suspended algae
- Algal thickness rating
- Abundance of algae at collection site
- Categorical % cover of algae at the collection site
- Presence of identifiable algal taxa

Can visible slime growths or nuisance aquatic growths be defined quantitatively?

- Measures can then be linked by stressor-response relationship
 - ▣ Measurements of algae:
 - Concentration of benthic & sestonic chlorophyll a
 - Organic biomass of benthic algal sample
 - Assemblage metrics (not yet calculated)
 - ▣ Measurements of Water Quality:
 - TN & TP
 - Orthophosphate, NO₃NO₂, TKN, turbidity, N:P

Are nutrients associated with these growths in a stressor-response context?

- A conceptual model is always a good idea

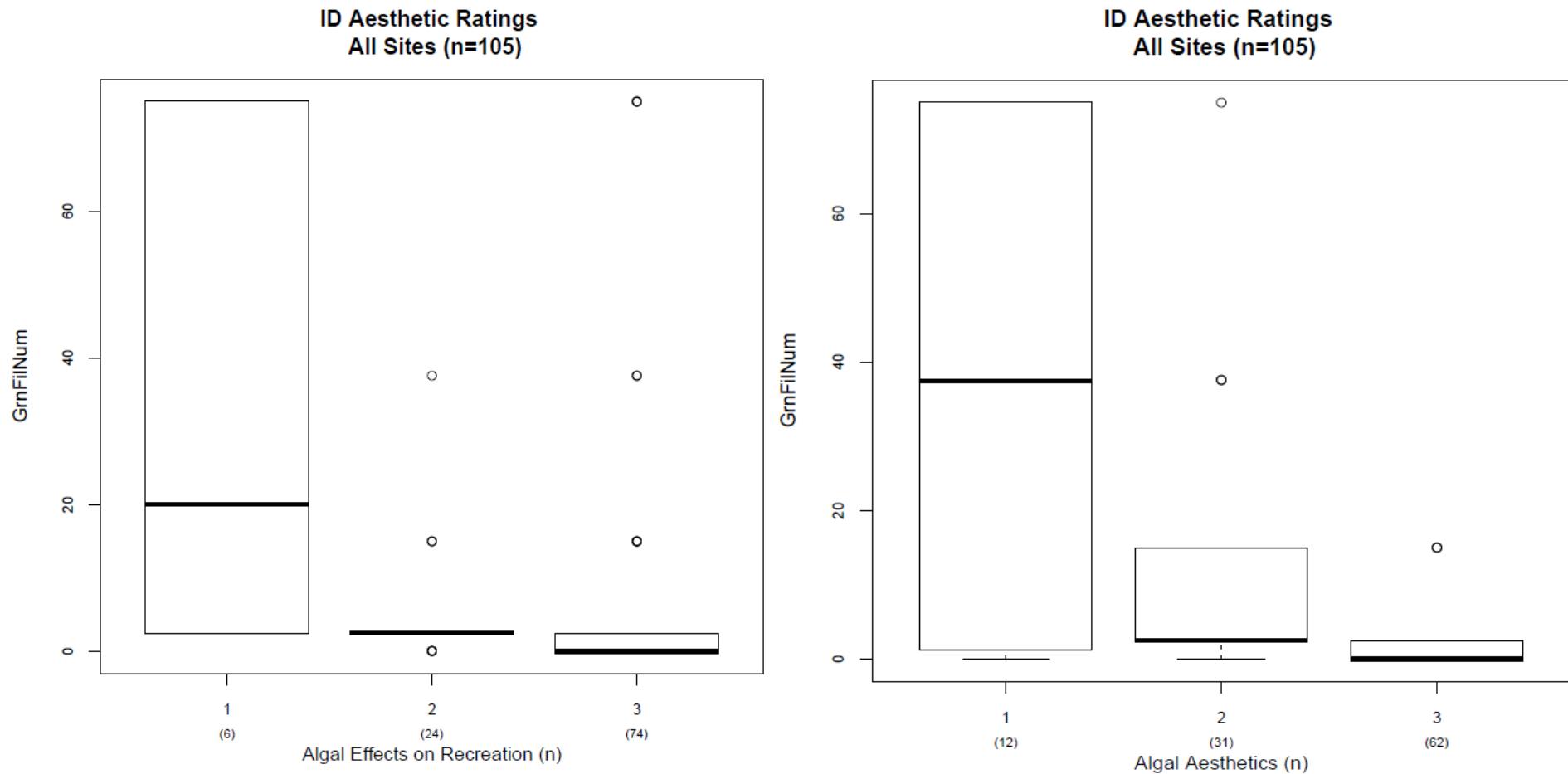


Are nutrients associated with these growths in a stressor-response context?

Idaho Dataset

- Six of 105 sites were rated as having “major” algal effects on recreational uses
 - ▣ Another 24 had “minor” effects
- Plots of measurements in relation to these ratings showed some variables as indicators
 - ▣ Green filamentous algae, floating mats and scum, and suspended algae were less common in sites with good algal aesthetics and no algal effects on recreation
 - ▣ On average, there were fewer macroalgae taxa and less frequent occurrence of *Spirogyra* in sites with no algal effects on recreation and good aesthetic ratings
- A quantitative measure of algal effects would be more defensible than a rating for defining nuisance algae

Are nutrients associated with these growths in a stressor-response context?

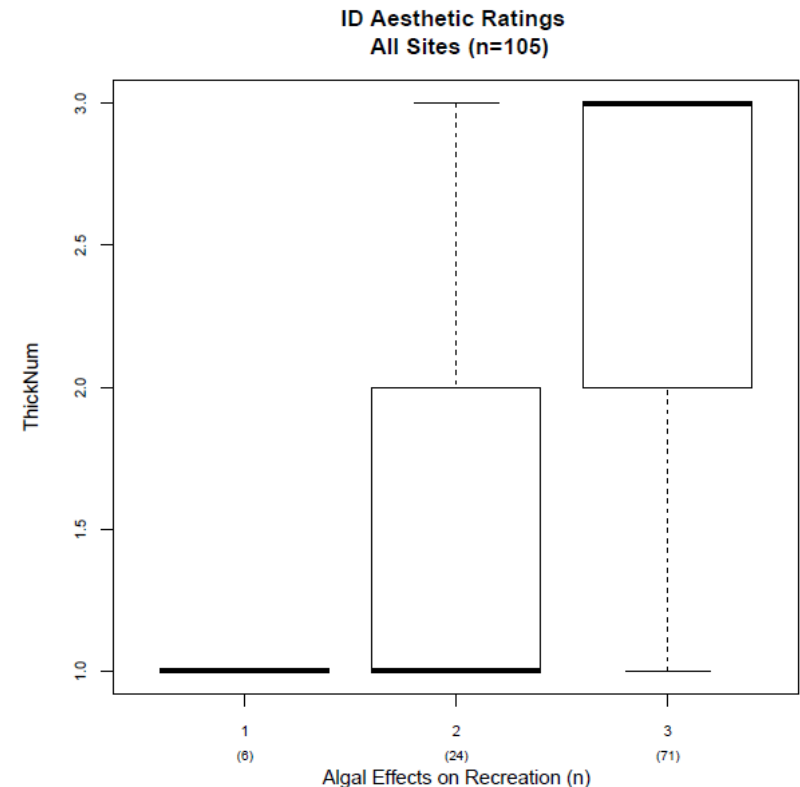


Green filamentous algae is related to algal effects and aesthetics

Are nutrients associated with these growths in a stressor-response context?

Variable	Aesthetics r	Algal Effects r
Green Filamentous	-0.47***	-0.30**
Floating Scum	-0.41***	-0.50***
Thickness	0.26**	0.67***
Algal Density	-0.55***	-0.01

Algal density measured on substrates was not related to algal effects – probably because substrates were sandy in streams with “major” effects.



Thickness was “visible” (rated 1) in all sites rated with algal effects on recreation

Major Algal Effects on Recreation



24479115 - Middle transect looking toward the left bank

Major Algal Effects on Recreation



24479159 - Middle transect looking toward the right bank

Major Algal Effects on Recreation



24479597 – Floating mats near middle transect

Major Algal Effects on Recreation



24482321 – Near middle transect looking at floating mats and a submerged tire

Major Algal Effects on Recreation



24491476 - Upper transect looking downstream

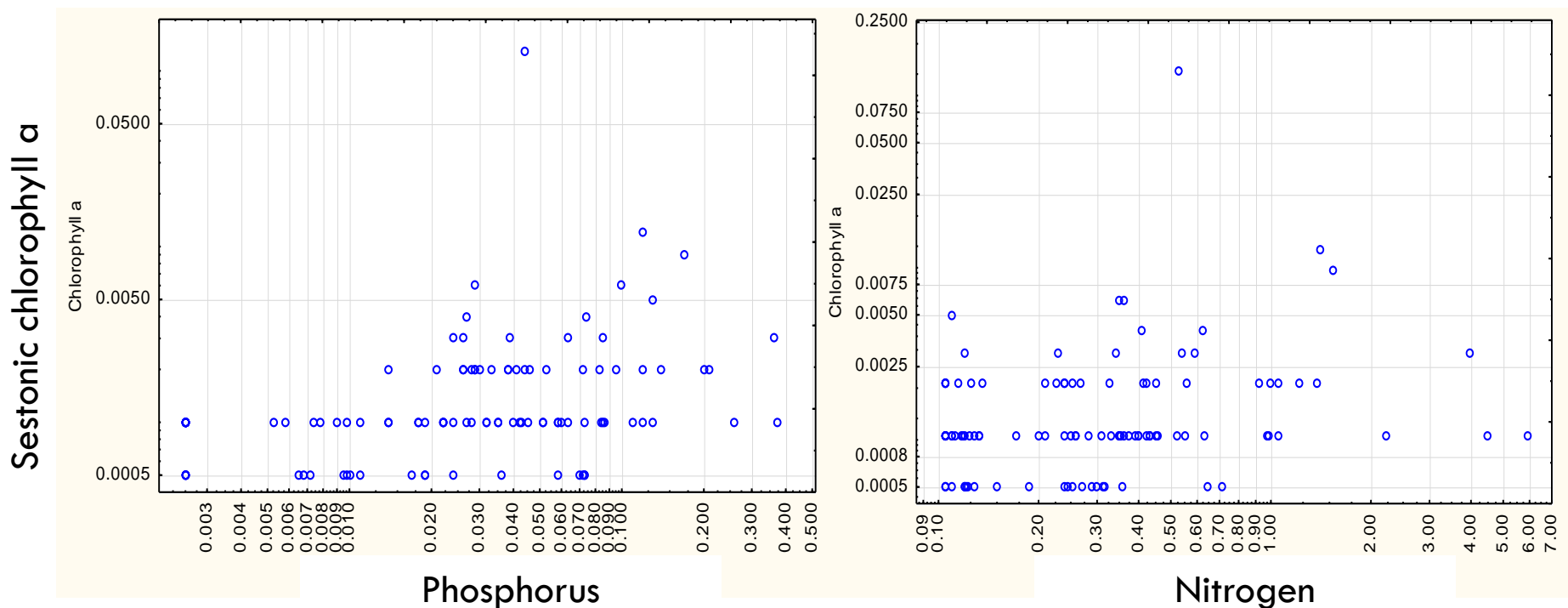
Major Algal Effects on Recreation



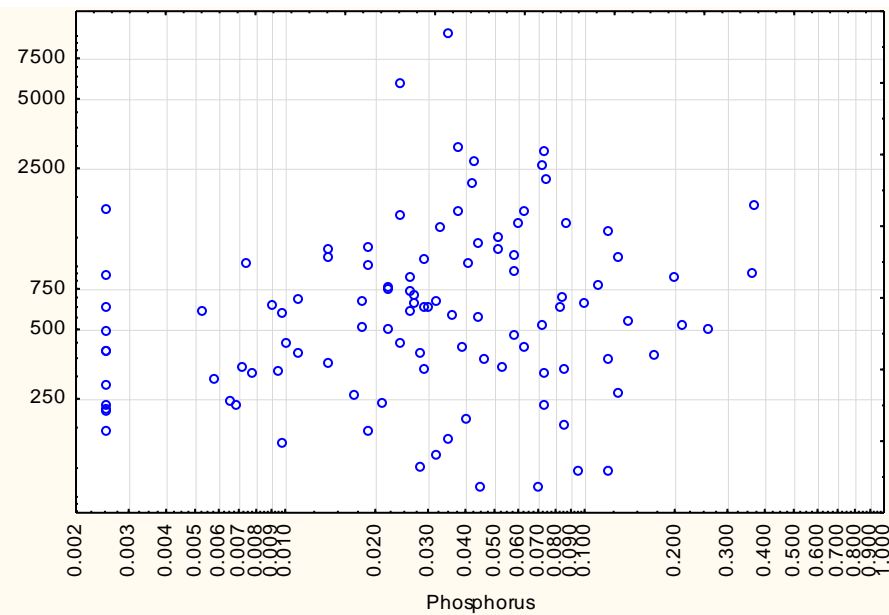
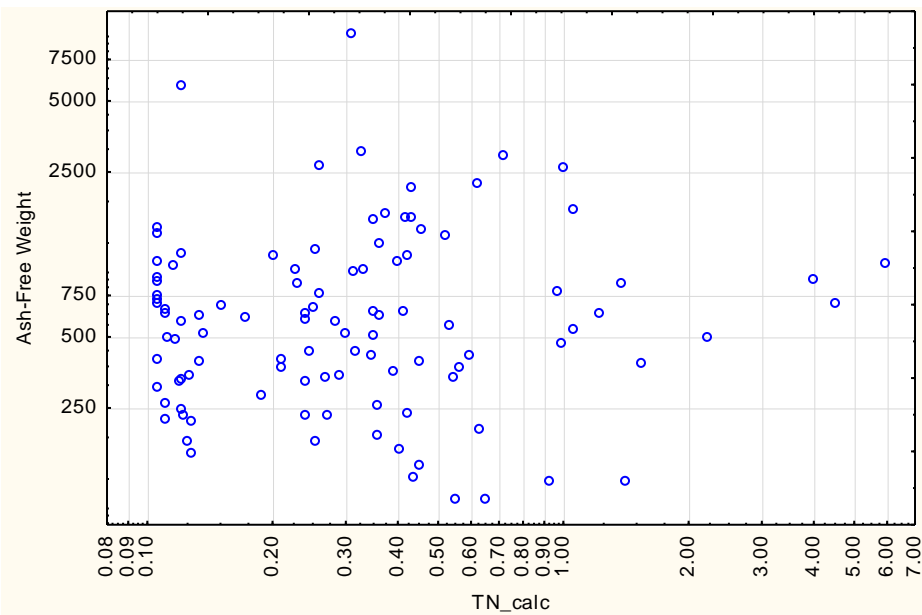
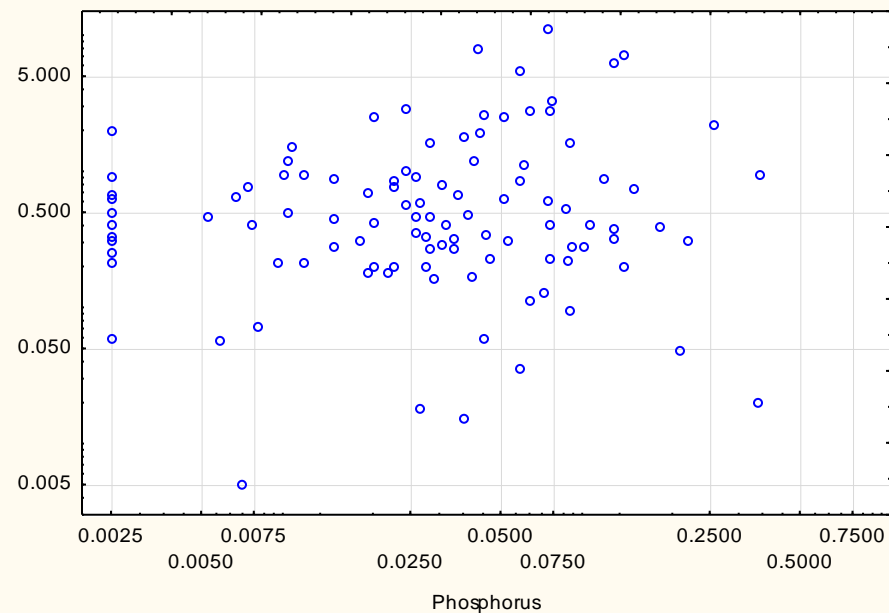
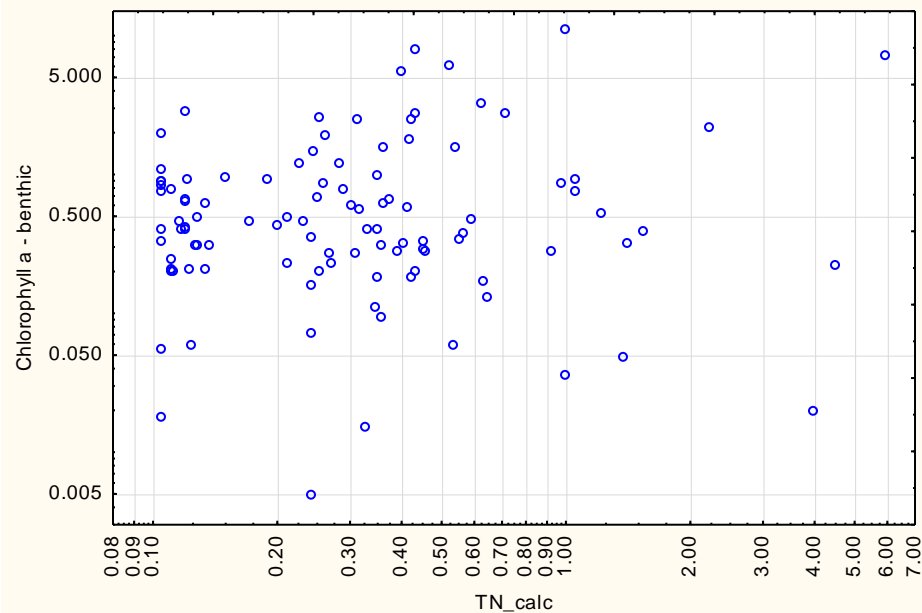
24562063 – Thick, slimy, layer of algae on a cobble

Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

- What is the stressor-response relationship with nutrients and algae?



Benthic Chlorophyll a



Benthic Algal Biomass (AFDM)

Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

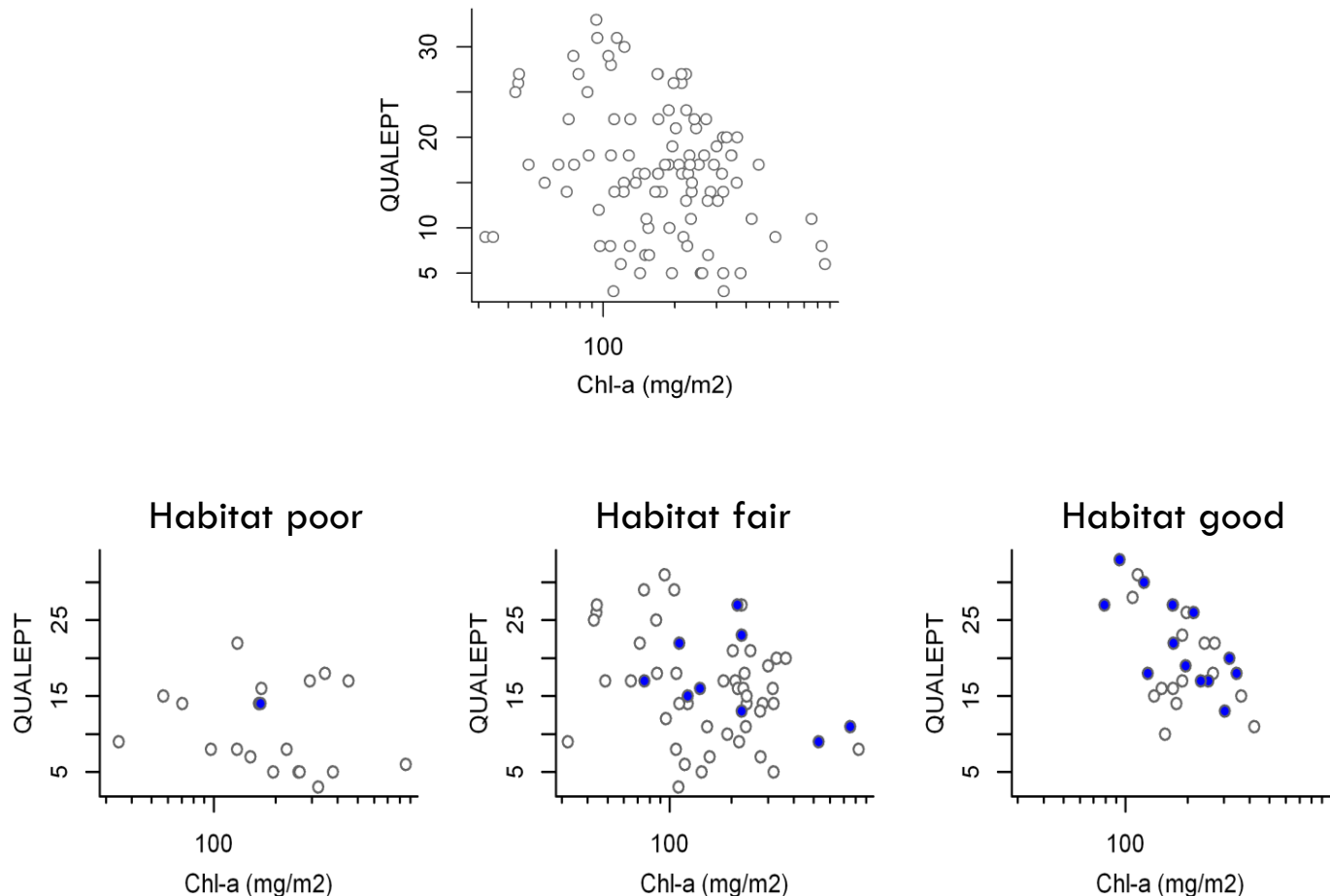


Confounding (or Classification) Factors

- Would we see stronger nutrient-algal relationships if we factored out multiple stressors or natural variables?
- Does this strengthen the assertion that nutrients are causes for responses?
- How do we factor them out?
 - ▣ Site classification
 - ▣ Partial correlation
 - ▣ Adjustment to regressions
 - ▣ Propensity scores

Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

Confounding (or Classification) Factors (example)

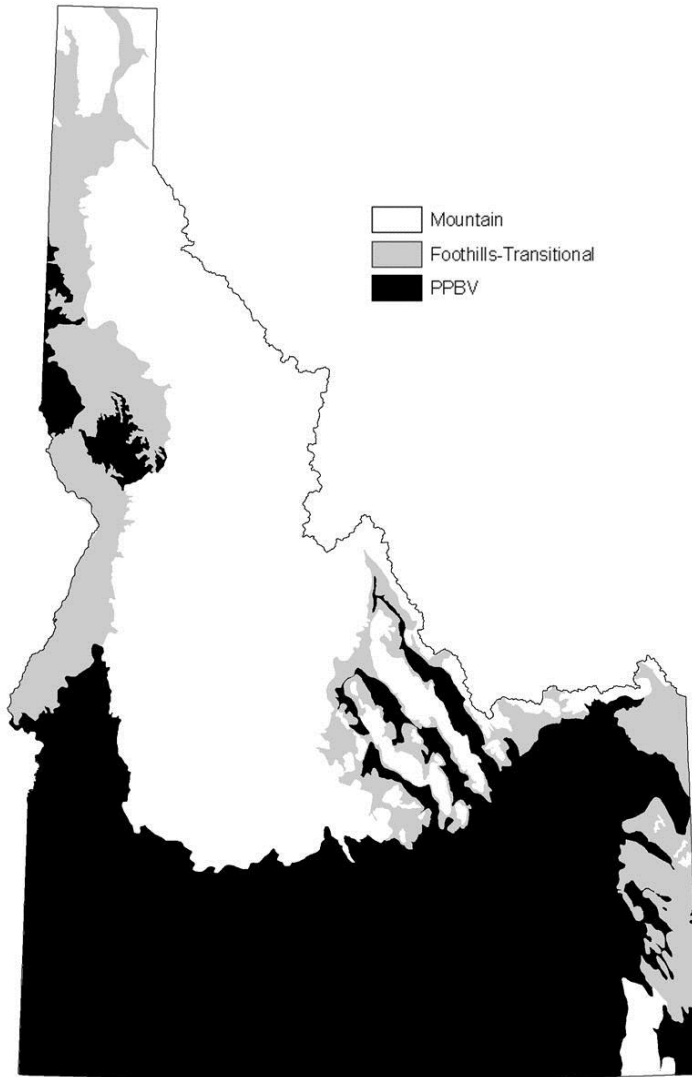


Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

Confounding (or Classification) Factors

- Natural (with some human influence?)
 - ▣ Light (canopy cover)
 - ▣ Temperature
 - ▣ pH
 - ▣ Ecoregion (and other GIS variables)
 - ▣ Flow
 - ▣ Gradient/scouring
 - ▣ Substrate
- Stressors
 - ▣ Conductivity
 - ▣ Land use (and other GIS variables)

Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?



Site Classification

- Start with an existing geographic scheme
 - ▣ Developed for biological assessment
 - ▣ 3 stream classes;
 - Mountains,
 - Foothills, and
 - Plains/Plateaus/Broad Valleys (PPBV).

Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

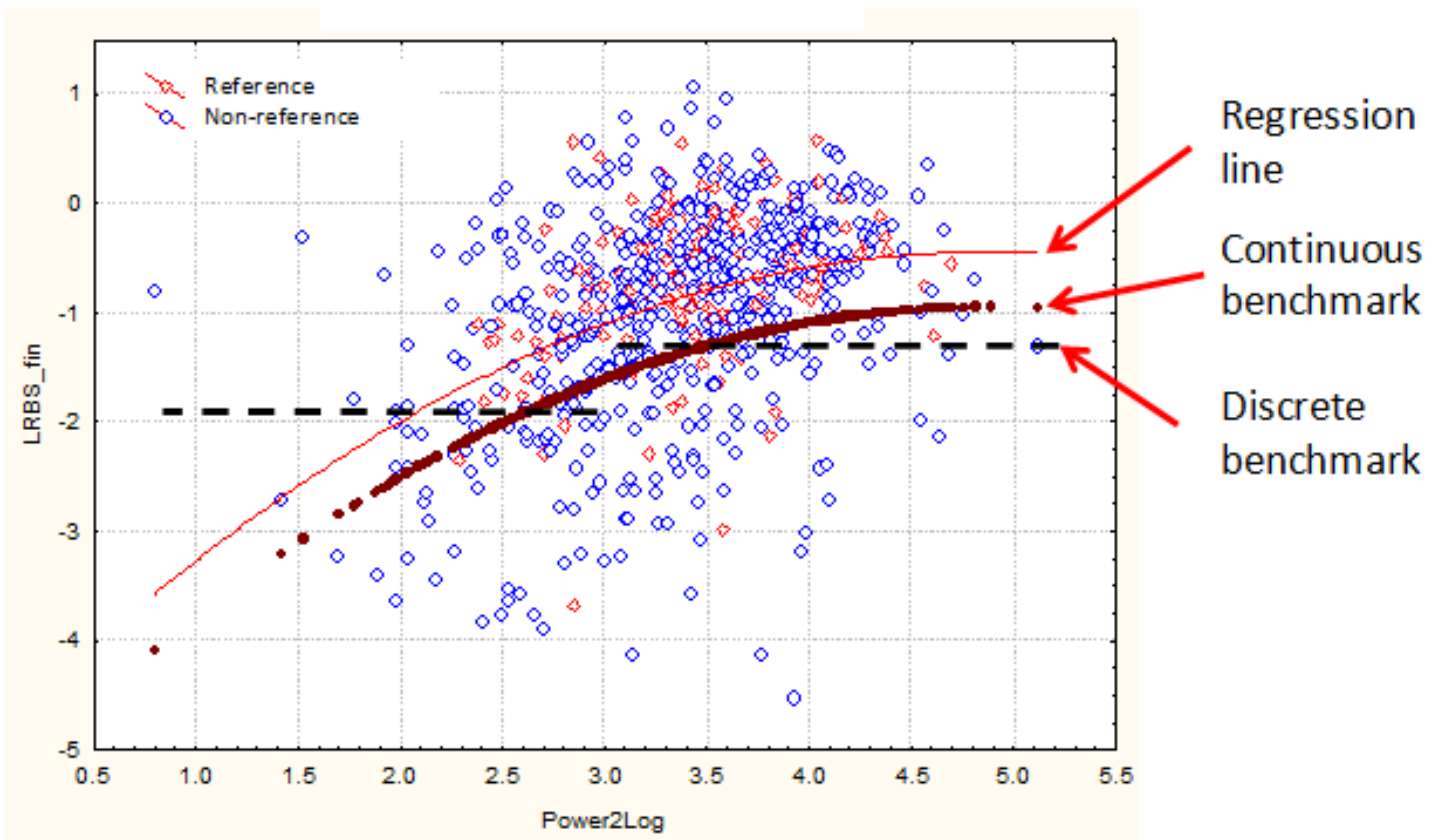
Partial Correlations

- Partial correlation controls for multiple factors, so that direct relationships can be explored
- This can help identify the nutrient-response relationships that exist despite underlying factors
- Partial correlations that are significant in all sites but that are not in site classes suggest that site classes control for some factors

Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

Adjustment to Regressions

- If a relationship is recognized, the adjustment to the natural

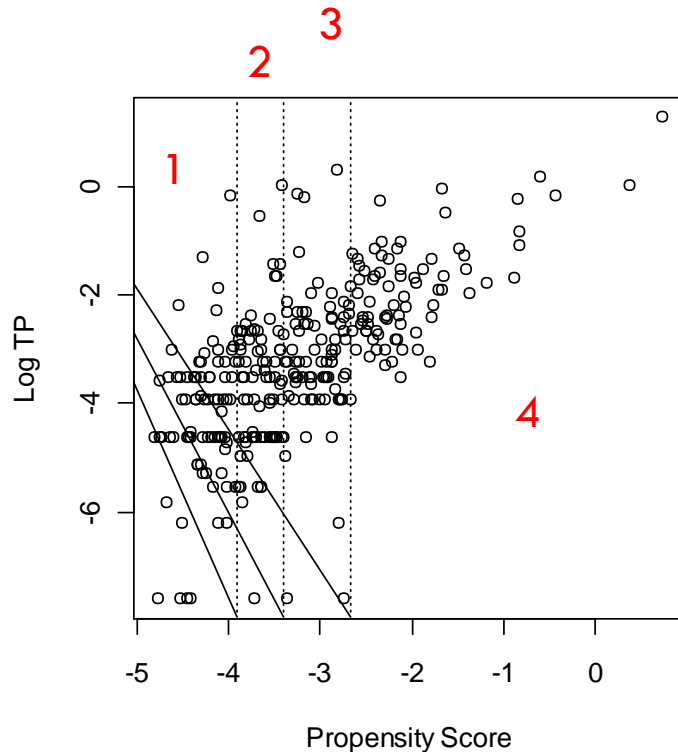


Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

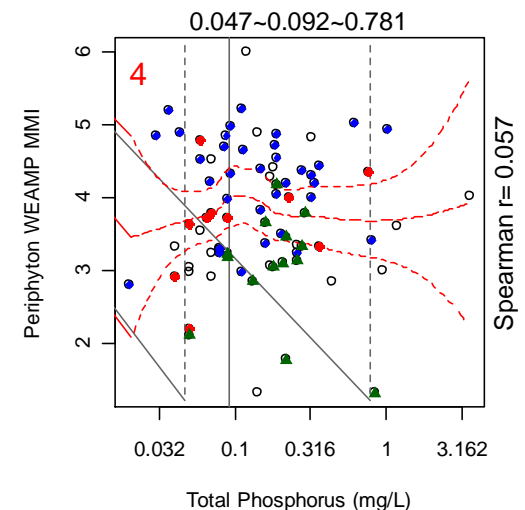
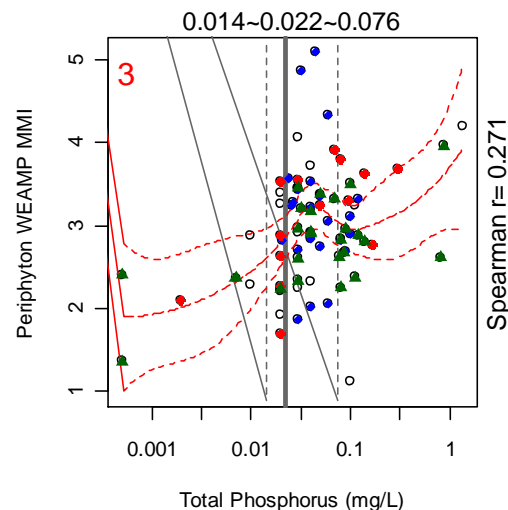
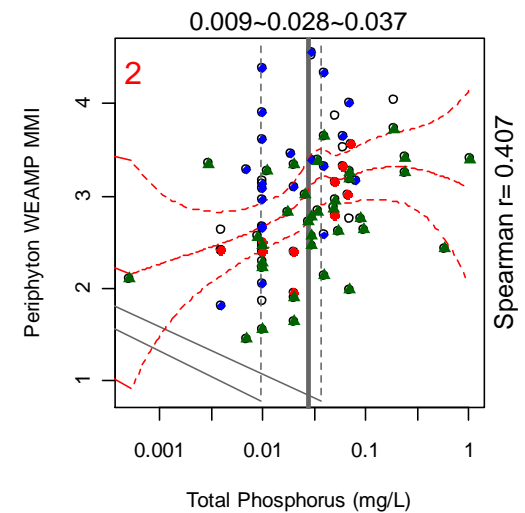
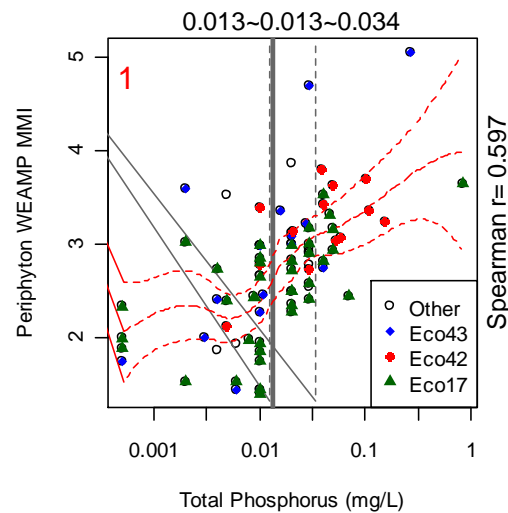
Propensity Scores

- Accounts for background effects of multiple co-varying stressors before indicating independent effects of nutrients
- A propensity function is the conditional probability of a multivariate treatment (e.g., nutrient concentrations), given values of known covariates
- Nutrient response relationships are analyzed within strata of the propensity scores

Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?



In Montana, we showed that TP had an effect on a periphyton MMI at values $<0.03\text{mg/L}$. Above that (4), other factors were as effective on periphyton.



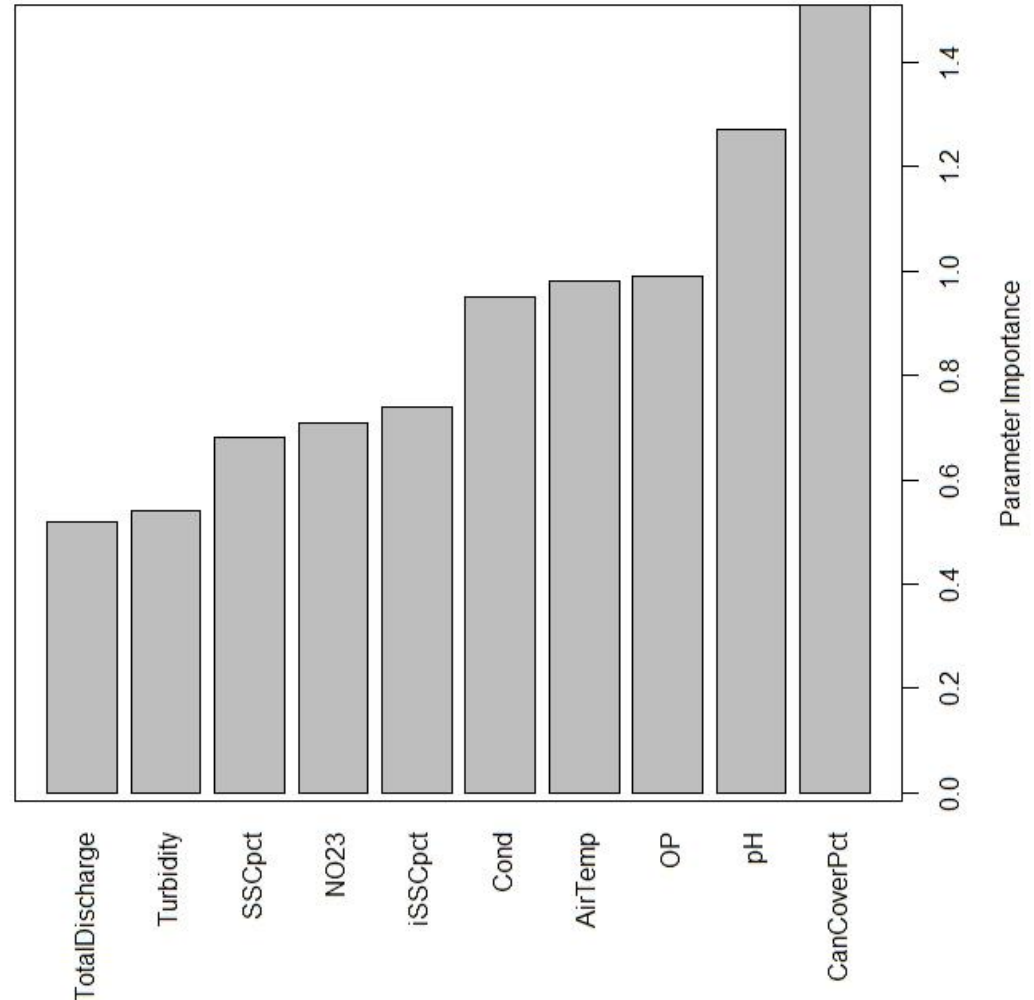
Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

- There are many more techniques to explore confounding/covariable effects
 - ▣ Multiple Regression
 - ▣ Classification and Regression Trees
 - ▣ Random Forests
 - ▣ Boosted Regression Trees
 - ▣ TREED models
- Once classes are identified, simplified models within each class may be easier to use, explain, and derive numeric criteria.

Can nutrient benchmarks be established to protect against unwanted visible slime growths or nuisance aquatic growths?

For Example

- Random Forests
 - Predicting benthic chlorophyll a
 - Variable importance
 - How frequently this predictor was chosen
 - Guides classification or model building



Duration, Frequency & Magnitude?

- Magnitude:
 - ▣ All our data are grab samples
 - ▣ Magnitude-based thresholds are appropriate
- Frequency:
 - ▣ In part, can be informed from precision, only in part
 - ▣ If we know the precision of replicate measures, can estimate the confidence in a single sample (and then multiples)
 - ▣ Ecological resilience/resistance concepts ought to apply too
- Duration:
 - ▣ Generally match underlying data;
 - ▣ But interesting thought exercise too....

Data Gaps: Duration, Frequency & Magnitude

- All our data are grab samples
 - ▣ Magnitude-based thresholds are appropriate
- Frequency may be derived from precision
 - ▣ If we know the precision of replicate measures, we can estimate the confidence in a single sample (and then multiples)
- Duration: ??

General Discussion